



Original Investigation | Oncology

# Assessment of Nuclear ZEB2 as a Biomarker for Colorectal Cancer Outcome and TNM Risk Stratification

Rahul Sreekumar, MD; Scott Harris, PhD; Karwan Moutasim, MD, PhD; Ricardo DeMateos, PhD; Ashish Patel, BM; Katherine Emo, BM; Sophie White, BM; Tamer Yagci, MD, PhD; Eugene Tulchinsky, PhD; Gareth Thomas, MD, PhD; John N. Primrose, MD; A. Emre Sayan, PhD; Alex H. Mirnezami, MD, PhD

## Abstract

**IMPORTANCE** At present, patients with colorectal cancer (CRC) are risk stratified using TNM histologic features. More recently, an association between a mesenchymal phenotype and a high risk of disease recurrence and micrometastases has been recognized.

**OBJECTIVE** To investigate the association of the epithelial to mesenchymal transition (EMT)-inducing transcription factor ZEB2 (zinc finger E box-binding homeobox 2), survival outcomes, and the efficacy of ZEB2 as a biomarker when added as refinement to TNM staging after curative intent surgery for CRC.

**DESIGN, SETTING, AND PARTICIPANTS** ZEB2 expression was assessed using a previously validated scoring system as part of a prospective, observational, masked diagnostic study from January 1, 2008, to December 31, 2013. Data were prospectively collected and analyzed for association with oncologic outcomes from January 1, 2017, to December 31, 2018. An initial test cohort from an academic university medical center of 126 consecutive patients with CRC and, subsequently, an independent validation cohort of 210 patients were examined. ZEB2 positivity was scored by 2 independent, masked pathologists. External validity was tested using an open access gene expression portal. Nomograms were developed with or without ZEB2.

**MAIN OUTCOMES AND MEASURES** Systemic and local recurrence of CRC.

**RESULTS** The test cohort consisted of 126 consecutive patients (mean [SD] age, 72.7 [11.7] years; 61 [48.4%] male) and the validation cohort of 210 patients (mean [SD] age, 72.0 [10.6] years; 111 [52.9%] male). A total of 52 tumors (41.3%) in the test cohort and 104 (49.5%) in the validation cohort were scored nuclear ZEB2 positive. Survival analysis by the log-rank test found that ZEB2 expression was associated with a significant reduction in overall survival and disease-free survival in both cohorts. Cox proportional hazards regression analysis highlighted ZEB2 as an independent biomarker of shorter overall survival and disease-free survival. Analysis of node-negative disease (n = 222) identified ZEB2 as an independent biomarker of early recurrence and reduced survival. External validation confirmed these findings. Addition of ZEB2 expression to nomograms composed of conventional TNM risk factors improved the ability to identify patients at high risk of recurrence demonstrated by the improvement in concordance index in both test (0.73 to 0.77) and validation (0.82 to 0.87) cohorts.

**CONCLUSIONS AND RELEVANCE** The findings suggest that expression of ZEB2 is associated with poor oncologic outcome and distant recurrence. The study also found that the addition of ZEB2 to existing TNM classification improved the ability to stratify patients for risk of recurrence. The results

(continued)

## Key Points

**Question** What is the clinical utility of the epithelial to mesenchymal transition program in assessing outcome in patients with colorectal cancer?

**Findings** In this diagnostic study with validation in 2 independent cohorts of 336 individuals, expression of the epithelial to mesenchymal transition inducing transcription factor ZEB2 was associated with poor oncologic outcomes and improved ability to risk stratify patients when added to the TNM classification, potentially affecting 14 of 100 patients.

**Meaning** The findings suggest that epithelial to mesenchymal transcription factor ZEB2 is a biomarker with clinical potential and may improve TNM risk stratification and guide treatment in colorectal cancer.

+ [Invited Commentary](#)

+ [Supplemental content](#)

Author affiliations and article information are listed at the end of this article.

**Open Access.** This is an open access article distributed under the terms of the CC-BY License.

Abstract (continued)

of this study suggest that addition of ZEB2 expression status to the TNM staging system improves the ability to stratify patients at high risk of recurrence.

JAMA Network Open. 2018;1(6):e183115. doi:10.1001/jamanetworkopen.2018.3115

## Introduction

Colorectal cancer (CRC) is the most common gastrointestinal cancer in Western countries.<sup>1</sup> Surgery remains the mainstay of treatment; however, metachronous systemic and local recurrence from occult micrometastatic spread is common and the principal cause of mortality.<sup>2</sup> One method to reduce recurrence risk is by modern combination adjuvant chemotherapy. Nevertheless, decision-making on application of adjuvant chemotherapy is challenging and inexact. As a result, many patients are overtreated with adjuvant chemotherapy and exposed to its detrimental effects, whereas others at risk of early-stage disease are not offered adjuvant chemotherapy but develop recurrence from undertreated micrometastases. These observations underscore the imprecise methods for staging, identification of micrometastases, and selection for adjuvant chemotherapy and highlight the critical need for better markers to identify occult spread.<sup>3</sup>

For decades, the TNM staging system, based on depth of tumor invasion and lymphatic node or distant organ spread, has been used to stratify patient risk and determine the need for adjuvant chemotherapy.<sup>4</sup> More recently, there has been increasing acknowledgment of the limitations of the TNM staging system and its failure to incorporate novel biological markers that may better inform disease trajectory.<sup>5</sup> Consequently, refinement of the TNM staging system through addition of validated molecular biomarkers of micrometastatic spread is an attractive concept and may improve recognition of patients whose conditions are currently erroneously understaged or overstaged.

Epithelial to mesenchymal transition (EMT) is a conserved epigenetic program that generates motile invasive mesenchymal cells from epithelial sheets under physiologic conditions.<sup>6</sup> In cancer, aberrant activation of this molecular program endows cancer cells with enhanced metastatic properties. Epithelial to mesenchymal transition is activated by signaling pathways that upregulate expression of EMT-inducing transcription factors.<sup>7</sup> In CRC, numerous in vitro and preclinical studies<sup>8-10</sup> have found that EMT leads to increased metastatic capacity and apoptosis resistance to commonly used chemotherapeutic agents. An association between EMT and poor oncologic outcomes has also been reported in other solid organ malignant tumors.<sup>9,11-13</sup> More recently, high-profile independent studies<sup>14-17</sup> have conducted molecular profiling of CRC, and although they have differed in the number of subtypes identified, tumors with an EMT-mesenchymal signature are consistently associated with poor outcomes. Nevertheless, the presence of a mesenchymal phenotype is not currently considered when stratifying recurrence risk or deciding on application of adjuvant chemotherapy.

Transcription factors that belong to the zinc finger E box-binding homeobox (ZEB) families are potent inducers of EMT<sup>7</sup> and comprise 2 genes, *ZEB1* (OMIM 189909) and *ZEB2* (OMIM 605802).<sup>18</sup> Aberrant expression of these genes is associated with poor outcomes in a number of malignant tumors.<sup>9,11,19</sup> To date, their potential contribution as a biomarker to improve the ability of TNM staging in identifying patients at high recurrence risk after curative surgery for CRC has not been investigated. This study investigates the association of the EMT-inducing transcription factor ZEB2, survival outcomes, and the efficacy of ZEB2 as a biomarker when added as refinement to TNM staging after curative intent surgery for CRC.

## Methods

### Setting and Participants

All patients were recruited as part of an ongoing, prospective UK National Institute of Health Research Clinical Research Network study that is investigating the molecular pathology of CRC and was designed to identify novel biomarkers. Details from this ongoing study have been described previously.<sup>20-23</sup> Study oversight and monitoring were performed at an independent clinical research organization. All patients provided written informed consent. All data were deidentified. After recruitment and surgery, tissue samples were deposited in a UK Human Tissue Act–approved tumor bank. Pathologic verification of diagnosis and staging was applied in accordance with the Association of Coloproctology of Great Britain and Ireland guidelines.<sup>24</sup> The South Central Ethics Committee (Bristol, United Kingdom) approved this study.

In this study, all patients were recruited from January 1, 2008, to December 31, 2013, after referral to a regional academic university medical center, and data were prospectively collected and analyzed from January 1, 2017, to December 31, 2018. Paraffin-embedded tissue specimens were retrieved for all patients in the current study. Exclusion criteria were evidence of a hereditary tumor, R1/R2 surgical resection, presence of multiple tumors, tumors with histologically identified extensive necrosis, and tumors with synchronous metastases at presentation. The database was queried for information that related to patient demographics, preoperative risk, imaging, surgery, pathologic features, postoperative management, and oncologic outcomes.

### Antibody Generation

The ZEB2 antibody used in this study has been described<sup>9</sup> and validated previously<sup>13,25</sup> and was used at a 1:750 concentration. The specificity of the antibody and applicability of the scoring criterion were previously validated and reported.<sup>25</sup>

### Immunohistochemical Analysis and Scoring Method

All immunohistochemical analyses were conducted at the histochemistry research unit at University Hospital Southampton, United Kingdom, using an automated immunostaining device (Autostainer XL, Leica Biosystems). Because ZEB2 is a nuclear transcription factor, all stained sections were assessed for the presence of nuclear ZEB2 expression in neoplastic and normal tissue. Mesenchymal cells, such as fibroblasts or lymphocytes, in the tissue naturally express ZEB2 and served as a positive control.<sup>9,13,25</sup> Two independent, masked pathologists (G.T. and K.M.) scored the sections as ZEB2 positive or negative using previously established scoring criteria.<sup>25</sup> When a disparity in scores was noted, slides were rereviewed to reach a consensus. ZEB2 expression was correlated with oncologic outcomes to evaluate its role as a biomarker. All results have been reported in line with reporting standards for biomarker development proposed by the Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) and the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guideline (eTables 1 and 2 in the [Supplement](#)).<sup>26</sup>

### Survival Analyses, Power Calculation, and Open Access Gene Expression Analysis

SPSS statistical software, version 22 (IBM Corp) was used to undertake survival analysis and examine the association between clinical and pathologic features and ZEB2 expression. Pathologists masked to patient outcome performed all biomarker scoring. Primary study end points of overall survival (OS) and disease-free survival (DFS) were defined as the time from date of primary resection to the date of death for OS or recurrence for DFS. Patient outcomes are represented as Kaplan-Meier survival curves and differences in survival outcomes calculated using the log-rank test. Multivariable analysis using a Cox proportional hazards regression model was used to investigate the prognostic value of ZEB2 in a model that encompassed conventional pathologic risk factors, and hazard ratio (HR) tables were used to compare differences. The association of ZEB2 with clinical and pathologic factors was

determined using a  $\chi^2$  or Fisher exact test as appropriate. Power calculation was performed using nQuery statistical software (Statistical Solutions Ltd). The test cohort for ZEB2 biomarker evaluation consisted of 126 consecutive patients. On the basis of the test cohort, a power calculation identified a minimum sample size of 180 as a requirement to achieve 80% power using a 2-sided test and a significance of 5%, assuming an HR of 2.0. The final figure recruited for the validation cohort was 210 consecutive patients. Additional details of settings and participants are reported in eTable 3 in the [Supplement](#). The reported *P* values are 2-sided, and statistical significance was set to *P* < .05 for all tests in the study. The HRs are presented with 95% CIs. External validity was evaluated using the open access gene expression portal PROGeneV2,<sup>27</sup> which is compiled from Gene Expression Omnibus, the European Bioinformatics Institute's ArrayExpress, and The Cancer Genome Atlas. Messenger RNA (mRNA) expression profiles were queried for ZEB2 and disease relapse. Kaplan-Meier survival curves were generated and statistical significance was calculated by the log-rank test.

### Mathematical Modeling and Nomogram Generation

A binary logistic regression model composed of conventional pathologic risk factors with or without the addition of the ZEB2 score was used to construct nomograms in the test cohort. The validity of the model was investigated by applying the model to the validation cohort. The capacity to determine the risk of distant recurrence within 3 years of surgical resection was calculated before and after the addition of ZEB2 to TNM staging criterion, differentiation, and extramural vascular invasion. The capacity of ZEB2 to determine outcome when added to the nomograms is reported as a concordance index (C index), and the contribution of adding ZEB2 to the nomograms is reported as incremental area under the curve (iAUC) in accordance with American Joint Committee on Cancer guidelines.<sup>6</sup>

The C index estimates the probability of concordance between the estimated and observed outcomes in rank order and equivalent to the area under the receiver operating characteristic curve. The C index represents the ability of the model to discriminate between patients who developed distant recurrence and those who did not. The iAUC represents the improvement in the C index as a result of addition of the ZEB2 expression score. With an equal interest in sensitivity and specificity, the optimum thresholds were selected to generate 2 risk scores with and without ZEB2 expression status. Patients scoring equal to or above the threshold were classified as high risk and below the threshold as low risk. Kaplan-Meier and calibration plots were then generated to assess the clinical utility of the risk score. An identical method was followed during subset analysis of patients with node-negative disease. Risk factors used in the stages I to II model included the following: age, American Society of Anesthesiologists class, presence of bowel obstruction, presence of bowel perforation, T stage, differentiation, extramural vascular invasion, perineural invasion, and lymphatic invasion with or without ZEB2 score. The aim of constructing these nomograms was to investigate whether the addition of ZEB2, a mesenchymal cancer cell marker, increases the ability to identify patients at higher risk of early recurrence when used in conjunction with conventional pathologic or clinical risk factors.

## Results

### Patient Demographics

The test cohort consisted of 126 consecutive patients (mean [SD] age, 72.7 [11.7] years; 61 [48.4%] male) and the validation cohort of 210 patients (mean [SD] age, 72.0 [10.6] years; 111 [52.9%] male). Patient demographics for the cohorts are provided in eTable 3 in the [Supplement](#). In the test cohort, 89 patients (70.6%) had colonic carcinomas and 33 (26.2%) had rectal carcinomas. Disease staging was 1 or 2 in 77 patients, and 48 had lymph node metastasis (median lymph node sample of 14). In the validation cohort, 159 patients (75.7%) had colonic carcinomas and 51 (24.3%) had rectal carcinomas (eTable 3 in the [Supplement](#)). A total of 67 patients (31.9%) had lymph node metastasis

(median node harvest of 15). A total of 92% of all rectal cancers were resected in the mesorectal plane.

### Association of Nuclear ZEB2 With Early Recurrence and Reduced Survival in the Test Cohort

Immunohistochemical analysis was performed on the test cohort and ZEB2 expression and positivity recorded using a previously described scoring system.<sup>9</sup> The REMARK and TRIPOD biomarker standards for the present study are provided in eTables 1 and 2 in the [Supplement](#). Representative images of ZEB2 immunostaining are shown in [Figure 1](#). No ubiquitous expression of nuclear ZEB2 was detected in normal colonic epithelium ([Figure 1A and B](#)). Mesenchymal cells, such as fibroblasts or lymphocytes, naturally express ZEB2 and served as the positive control. Two pathologists (G.T. and K.M.) masked to groups and outcomes scored 52 of the 126 specimens (41.3%) as ZEB2 positive ([Figure 1C-H and eTable 3 in the Supplement](#)).

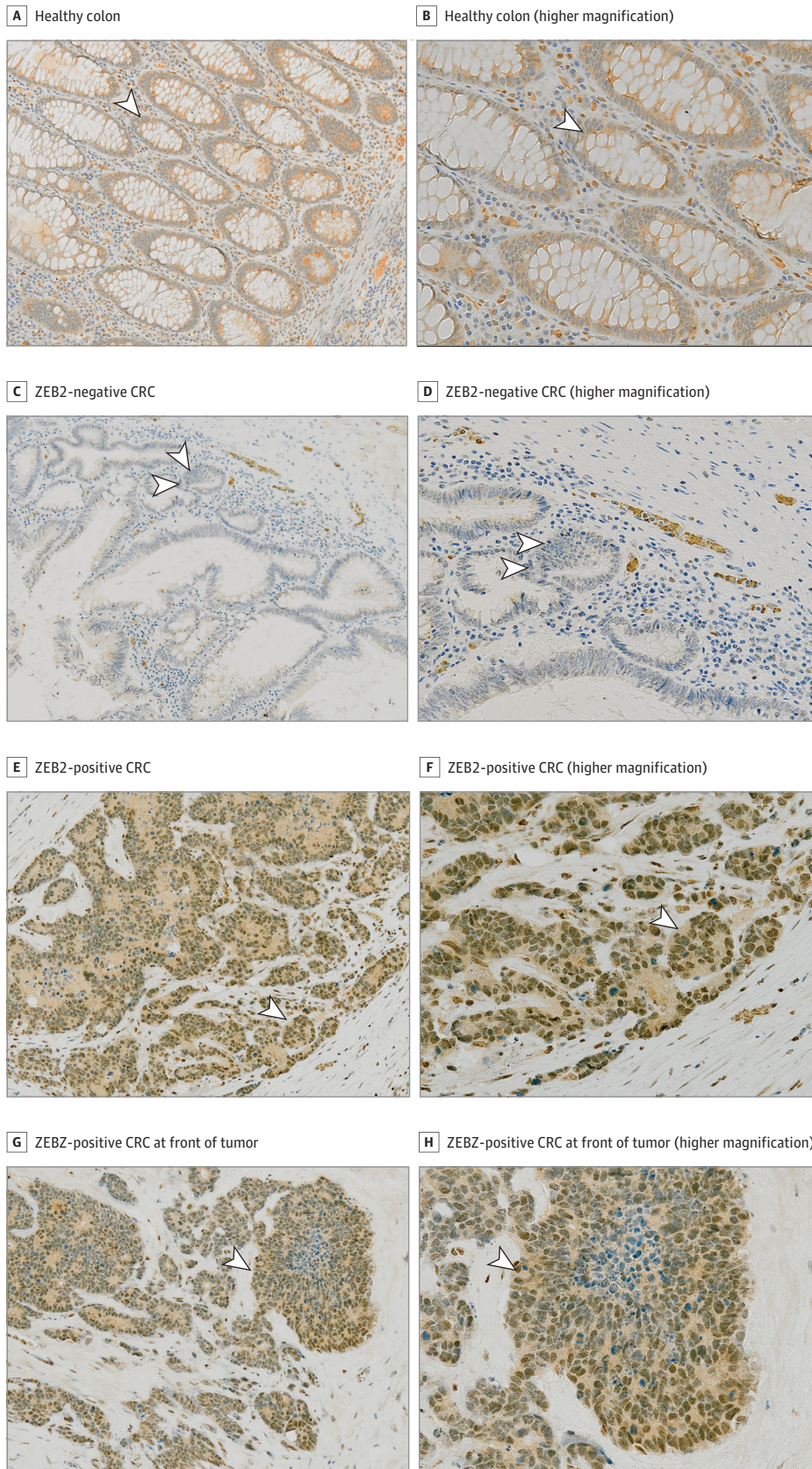
Clinical and pathologic evaluation revealed a potential association between ZEB2 positivity and lymph node metastasis and higher-stage disease ([eTable 4 in the Supplement](#)). No association was observed between ZEB2 expression and differentiation, age, sex, or T stage. Survival analyses by log-rank test highlighted increased recurrence ([Figure 2B](#)) and reduced survival ([Figure 2A](#)) in ZEB2-expressing patients. Mean OS among ZEB2-expressing patients was 43.8 months (95% CI, 34.4-49.7 months) compared with 60.4 months (95% CI, 50.0-65.8 months) among ZEB2-negative patients (log-rank test,  $P = .02$ ). Mean DFS among ZEB2-positive patients was 48.0 months (95% CI, 38.2-54.0 months) compared with 60.5 months (95% CI, 56.6-72.5 months) among ZEB2-negative patients (log-rank test,  $P = .005$ ). Multivariable analysis by Cox proportional hazards regression analysis highlighted ZEB2 as an independent biomarker of both OS (HR, 1.7; 95% CI, 1.1-2.8;  $P = .05$ ) and DFS (HR, 2.2; 95% CI, 1.2-4.2;  $P = .05$ ) ([eTable 5A in the Supplement](#)). Compared with ZEB2-negative patients, ZEB2-positive patients had a 1.7-fold increased risk of mortality and a more than 2-fold increased recurrence within 5 years of surgery.

### ZEB2 Analysis in Validation Cohort

Informed by the test cohort, a power calculation identified the need for a minimum of 180 patients in a validation cohort. The final validation cohort consisted of 210 consecutive patients to allow for exclusions. Patients with metastatic disease at diagnosis ( $n = 26$ ) were excluded. Assessment of ZEB2 immunoexpression revealed 104 of 210 tumors (49.5%) to be ZEB2 positive ([eTable 3 in the Supplement](#)). Clinical and pathologic evaluation mirrored associations observed in the test cohort ([eTable 4 in the Supplement](#)). Survival analysis by log-rank testing maintained consistency with results from the test cohort. A mean 6.9-month reduction in OS (ZEB2-negative tumors vs ZEB2-positive tumors: 47.0 [95% CI, 42.3-51.7] vs 53.9 [95% CI, 48.8-57.9] months; log-rank test,  $P = .02$ ) ([Figure 3A](#)) and a mean 13-month reduction (ZEB2-negative tumors vs ZEB2-positive tumors: 44.6 [95% CI, 38.9-50.0] vs 57.6 [95% CI, 53.1-61.3] months; log-rank test,  $P = .001$ ) ([Figure 3B](#)) in DFS were observed among ZEB2-positive compared with ZEB2-negative patients. Multivariable analysis verified ZEB2 as an independent biomarker of OS (HR, 1.4; 95% CI, 1.2-2.1;  $P = .05$ ) and DFS (HR, 3.0; 95% CI, 1.6-6.1;  $P = .001$ ) ([eTable 5B in the Supplement](#)). T stage, N stage, and ZEB2 positivity were identified as independent biomarkers of OS, whereas T stage, N-stage differentiation, and ZEB2 status were independent biomarkers of DFS. We next investigated external validity of our findings using the open access gene expression portal PROGgeneV2, which is compiled from Gene Expression Omnibus, the European Bioinformatics Institute's ArrayExpress, and The Cancer Genome Atlas. This search confirmed findings similar to our test and validation cohorts, with high ZEB2 expression associated with a greater risk of recurrence (HR, 2.1; 95% CI, 1.16-3.79;  $P = .02$ ) ([eFigure 1 in the Supplement](#)).



**Figure 1. Immunohistochemical Analysis of ZEB2 Expression in Healthy Colon and Colorectal Cancer (CRC) Specimens**



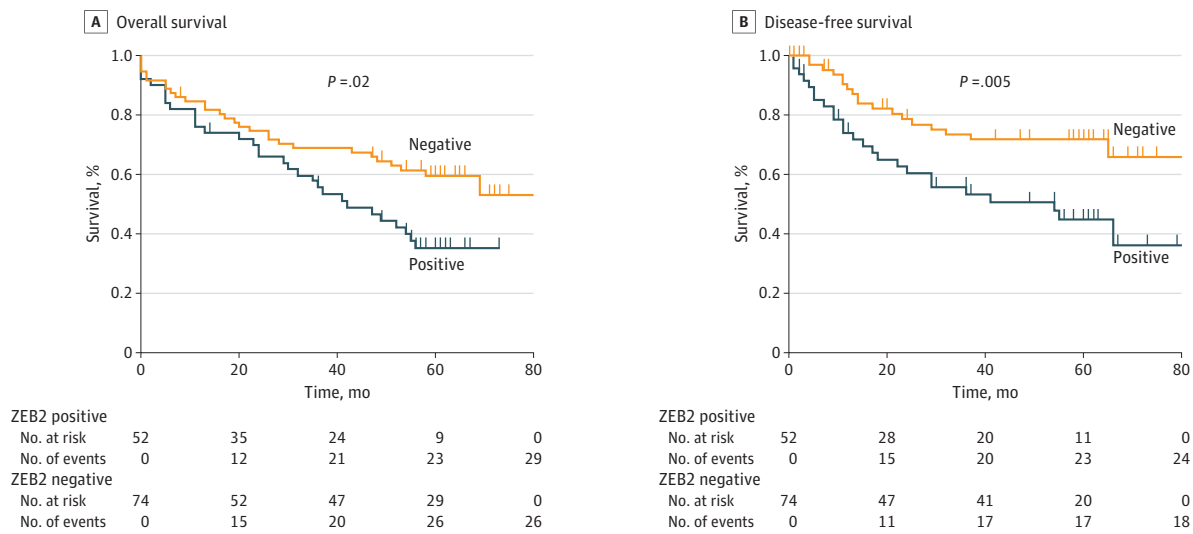
A and B, Healthy colonic epithelium. The arrowheads indicate cells with an absence of ZEB2 nuclear staining (A) and fibroblasts and immune cells with positive nuclear staining that served as internal positive controls (B). C and D, Representative example of a CRC specimen that stained negative for ZEB2. Arrowheads indicate cells with an absence of ZEB2 nuclear staining (C) and fibroblasts and immune cells that served as internal positive controls (D). E and F, Representative example of a CRC specimen that stained positive for ZEB2. Arrowheads indicate CRC specimen expressing nuclear ZEB2 in neoplastic cells (E) and staining of cells in the middle of the cluster of neoplastic cells and evidence of the specificity of the antibody (F). G and H, A CRC specimen that was ZEB2 positive at the invasive front of the tumor. Immunohistochemical analysis for ZEB2 (brown) was performed on all images, visualized by 3,3'-diaminobenzidine–positive chromogen, and counterstained with hematoxylin (blue) (original magnification  $\times 100$  [A, C, E, and G] and  $\times 200$  [B, D, F, and H]). Arrowheads indicate nuclear positivity.

**Association of Nuclear ZEB2 With Risk of Distant but Not Local Recurrence**

Several in vitro and in vivo studies<sup>9,28,29</sup> have reported that EMT results in an enhanced capacity to metastasize to distant organs. Accordingly, we evaluated study patients separately for distant recurrence as opposed to local recurrence and examined the association of ZEB2 in these contexts. Detailed definitions of distant recurrence and local recurrence are described in this study’s REMARK profile (eTable 1 in the Supplement).

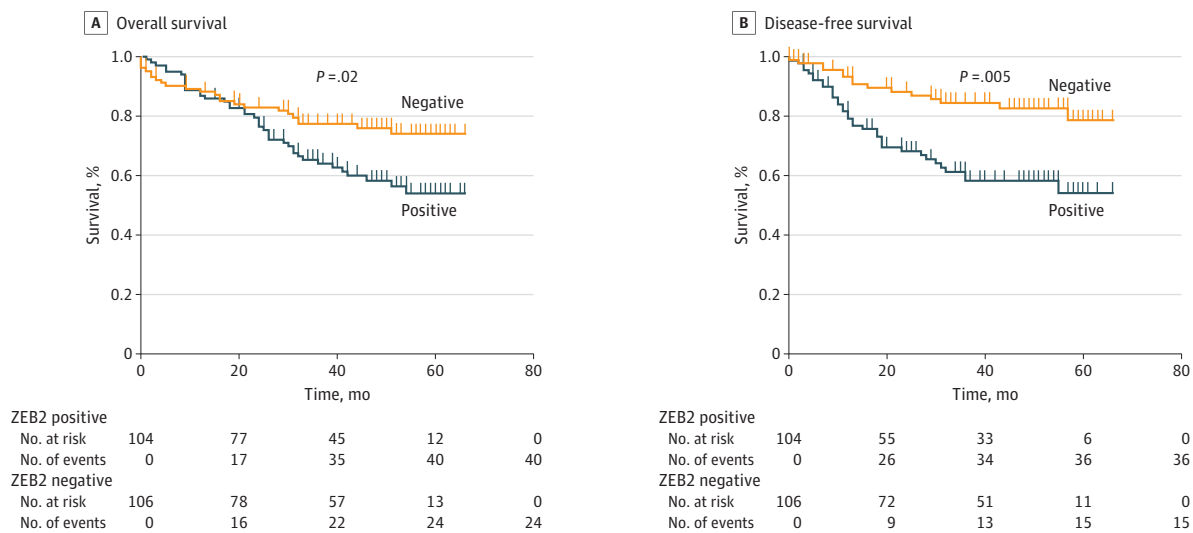
Survival analysis in the test and validation cohorts showed that ZEB2 positivity selectively determines distant recurrence but not local recurrence (eFigures 2 and 3 in the Supplement), consistent with enhanced motility and migration characteristics associated with EMT. Mean time to distant recurrence was significantly shorter in ZEB2-positive patients in both cohorts. An 18-month

**Figure 2. Association of ZEB2 With Early Recurrence and Reduced Survival in the Test Cohort**



Kaplan-Meier survival curves for the test cohort demonstrating differences in overall survival (A) and disease-free survival (B) when patients were stratified as ZEB2 negative or positive. *P* values were calculated using the log-rank test.

**Figure 3. Association of ZEB2 With Early Recurrence and Reduced Survival in the Validation Cohort**



Kaplan-Meier survival curves for validation cohort demonstrating differences in overall survival (A) and disease-free survival (B) when patients were stratified as ZEB2 negative or positive. *P* values were calculated using the log-rank test.

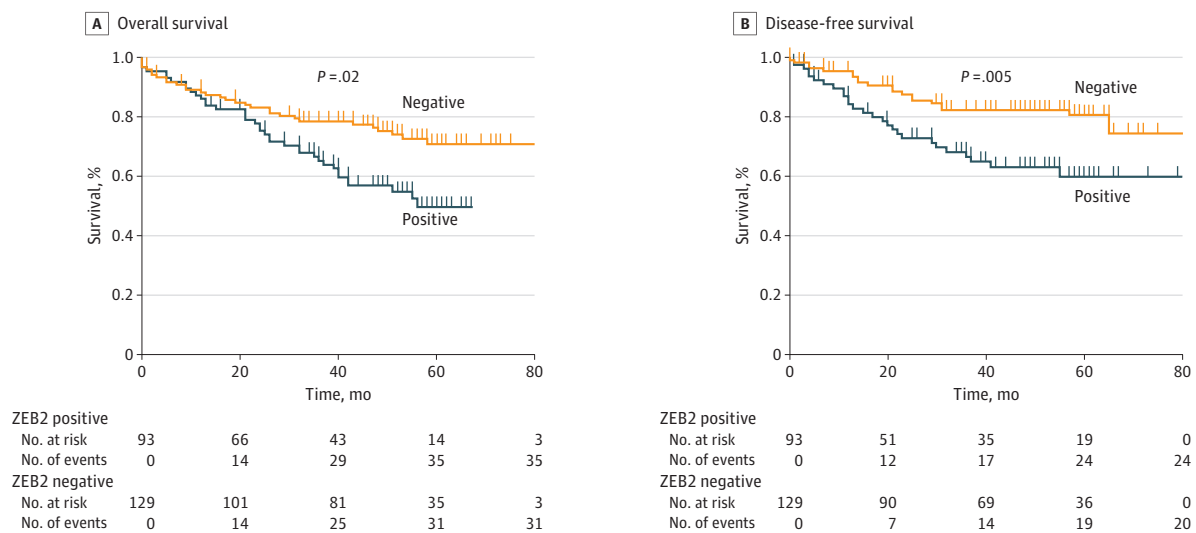
reduction in mean time to distant recurrence (49.6 [95% CI, 40.1-59.1] vs 67.6 [95% CI, 59.9-75.2] months; log-rank test,  $P = .01$ ) in the test cohort and a 11.4-month reduction (48.0 [95% CI, 42.6-53.5] vs 59.4 [55.6-63.1] months; log-rank test,  $P = .01$ ) in the validation cohort were observed. In contrast, no difference in local recurrence was observed. Multivariable analysis identified ZEB2 as an independent biomarker of distant recurrence in the test (HR, 1.8; 95% CI, 1.4-3.7;  $P = .05$ ) and validation (HR, 3.2; 95% CI, 1.6-6.6;  $P = .001$ ) cohorts.

**Identification of Stages I to II Disease in Patients at High Risk of Recurrence**

The benefit of administering adjuvant chemotherapy in stage III CRC is well recognized.<sup>30</sup> Selecting node-negative patients (stages I to II disease) who benefit from adjuvant chemotherapy using conventional assessment remains suboptimal, however.<sup>31</sup> Consequently, we next investigated whether ZEB2 expression in node-negative patients could identify patients at higher risk for recurrence. In total, 222 patients from the total study cohort were identified and cumulatively analyzed (demographics and clinicopathologic associations given in eTables 6 and 7 in the Supplement). ZEB2 expression in node-negative patients was associated with a significant reduction in OS and DFS. ZEB2-positive, node-negative tumors had a 19.4-month reduction in time to recurrence (mean, 55.5 [95% CI, 48.4-62.6] vs 74.9 [95% CI, 68.6-81.0] months; log-rank test,  $P = .001$ ) and 23.8-month decrease in OS (mean, 46.4 [95% CI, 41.0-51.5] vs 70.2 [95% CI, 64.2-76.3] months; log-rank test,  $P = .008$ ) (Figure 4). ZEB2 expression was associated with earlier distant recurrence (log-rank test,  $P = .01$ ) but not local recurrence (log-rank test,  $P = .12$ ).

An 18.2-month reduction (mean, 59.9 [95% CI, 53.1-66.8] vs 78.1 [95% CI, 72.3-83.9] months) in time to distant recurrence was observed in ZEB2-positive cases. Multivariable analysis incorporating known clinical and pathologic risk factors of stage I and II disease identified ZEB2 expression as an independent biomarker of distant recurrence (HR, 1.9; 95% CI, 1.6-6.6;  $P = .001$ ) and OS (HR, 1.9; 95% CI, 1.2-3.2;  $P = .009$ ) (eTable 8 in the Supplement). At 5 years, ZEB2-expressing patients were almost twice as likely to have experienced recurrence or death compared with ZEB2-negative patients. Further subgroup analyses to evaluate rectal and colon cancer separately found similar findings, and multivariable analysis identified ZEB2 as an independent biomarker (eFigure 4 in the Supplement).

Figure 4. Association of ZEB2 With Early Recurrence and Reduced Survival in Stage I/II Disease



Kaplan-Meier survival curves from analysis of patients with node-negative stage I/II disease, demonstrating differences in overall survival (A) and disease-free survival (B)

when patients were stratified as ZEB2 negative or positive. P values were calculated using the log-rank test.



## Addition of ZEB2 Expression to TNM Staging

Our findings confirm and validate an association between ZEB2 expression and increased incidence of distant recurrence, independent of stage. To investigate whether the addition of ZEB2 expression to conventional TNM risk factors is associated with an improvement in the ability to identify patients at higher risk of early recurrence (<3 years) after curative surgery, nomograms and mathematical modeling with or without ZEB2 expression scores were developed.

Nomograms and modeling were developed in the test cohort and applied to our validation set. In the test cohort, the C index to identify distant recurrence within 3 years of surgery using conventional histologic risk factors was 0.73 (95% CI, 0.62-0.84) and improved to 0.77 (95% CI, 0.66-0.87; iAUC, 0.04) with the addition of ZEB2 expression.

Application of the nomogram to the validation cohort (n = 210) showed excellent performance and ability to identify risk, highlighting an identical trend, with the C index improving from 0.82 (95% CI, 0.75-0.87) to 0.87 (95% CI, 0.80-0.94; iAUC, 0.05) (Figure 5A and B). Kaplan-Meier survival plots generated by applying the risk score (high risk,  $\geq 1.4$ ; low risk,  $< 1.4$ ) to the validation cohort revealed significantly greater separation (Figure 5C and D). Application of ZEB2 scoring resulted in improvement in sensitivity of 5.2% and specificity of 8.7%. These improvements would result in improved risk stratification in 14 of 100 patients assessed compared with use of the TNM staging system without inclusion of ZEB2-expression status. Calibration plots demonstrated good concordance between expected and observed outcome, and results of the Hosmer-Lemeshow test, used to detect differences in expected and observed events, were statistically insignificant ( $P = .84$ ). Taken together, these observations suggest that ZEB2 expression, if used in conjunction with conventional TNM-based histologic staging, is associated with an improvement in the ability to identify patients at increased risk of experiencing distant recurrence after surgical resection.

The inability to stratify risk in patients with node-negative CRC impedes the ability to identify patients who will derive maximum benefit from adjuvant chemotherapy. We next investigated whether the addition of ZEB2 expression to existing high-risk factors used to determine the need for adjuvant chemotherapy in patients with node-negative CRC would further aid detection of patients at increased risk for recurrence. A multivariable logistic regression model that contained conventional pathologic and clinical risks was constructed to identify independent variables. Clinical and pathologic variables considered and incorporated into the model are listed in eTable 8 in the Supplement. In node-negative disease, the C indices were 0.83 (95% CI, 0.70-0.95) for estimation of distant recurrence and 0.84 (95% CI, 0.78-0.90) for estimation of OS at 3 years. The C index improved from 0.80 to 0.83 with the addition of ZEB2, with improvements in sensitivity of 3% and specificity of 7%. Consequently, application of ZEB2 scoring would improve risk stratification in 10 of every 100 node-negative patients.

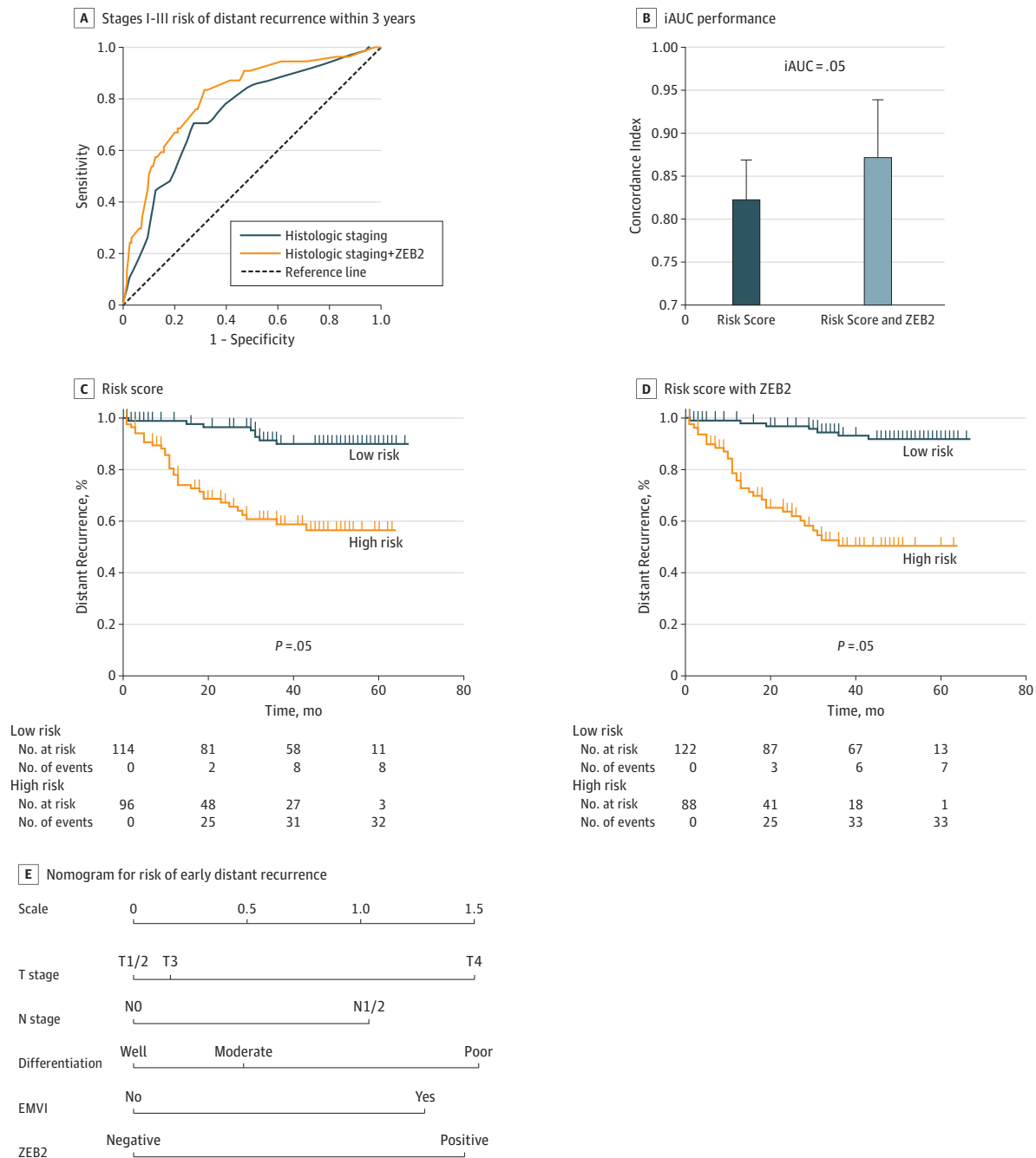
## Discussion

Currently, CRC is treated as a genetically homogeneous disease. There are no molecular biomarkers in routine clinical use to stratify risk of recurrence after surgical excision of the primary tumor, and only *kras* mutational status is approved as a negative determinant in patients treated with epidermal growth factor receptor inhibitors.<sup>32,33</sup> We demonstrate that ZEB2, an EMT-inducing transcription factor expressed in mesenchymal cancer cells, is associated with an increased risk of distant recurrence and reduced OS after curative surgery. Subset analysis in node-negative disease also highlighted ZEB2 as an independent biomarker. We next developed and internally validated, to our knowledge, the first nomogram for distant recurrence that incorporates this biomarker and found good estimation ability and superiority over conventional and currently used TNM risk factors.

Epithelial to mesenchymal transition is a critical event in cancer metastasis.<sup>6,34</sup> The association among EMT, downregulation of epithelial adhesion molecules,<sup>7</sup> modulation of the microenvironment to promote invasion,<sup>35</sup> and cytoskeletal alterations that increase motility<sup>7</sup> has long been known. Tissue-specific knockdown of critical EMT-inducing transcription factors (eg, ZEB1 in pancreatic

cancer) eliminates the ability of cancer cells to metastasize.<sup>28</sup> Furthermore, large-scale consortia of leading scientists have highlighted the importance of identifying tumors that express a mesenchymal phenotype.<sup>5,14-17</sup> Consequently, biomarkers that detect an EMT phenotype in the primary tumor may help identify those at risk of recurrence attributable to harboring occult micrometastases. E-cadherin

**Figure 5. Improvement in the Ability to Stratify Patients for Risk of Recurrence After Curative Surgery With the Addition of ZEB2 Expression Status to the TNM Staging System**



A and B, Receiver operating characteristic curves were generated with and without addition of ZEB2 to the TNM staging system to evaluate any improvement in incremental area under the curve (iAUC). B, Improvement in iAUC secondary to the addition of ZEB2 as a histogram. Error bars indicate 95% CI. C and D, Kaplan-Meier curves show increased separation and stratification when ZEB2 is added to the nomogram. Patients were

stratified as high risk or low risk based on scores generated from the nomogram. C, Results of the risk score alone. D, Results with the addition of ZEB2 scoring. P values were calculated using the log-rank test. E, Visual representation of the nomogram used to estimate risk of early distant recurrence. EMVI indicates extramural vascular invasion.

downregulation is a cardinal feature of EMT and has been extensively investigated as such a biomarker.<sup>12,36,37</sup> However, studies<sup>38,39</sup> that use E-cadherin have reported conflicting results.

Of the EMT transcription factors, ZEB proteins have been sparsely studied in CRC.<sup>40</sup> Furthermore, to date, no prospective or validated biomarker studies have been undertaken, making clinical translation challenging. Kahlert et al<sup>11</sup> previously reported cytoplasmic expression of ZEB2 at the invasive front of primary CRCs identified as having poor cancer-specific survival. However, nuclear expression, the ability to differentiate local recurrence from distant recurrence, or applicability in node-negative disease and ability to identify patients at high risk of distant recurrence were not considered. Our validated results emphasize the importance of nuclear positivity because ZEB proteins modulate gene expression through epigenetic regulation in the nucleus.<sup>41</sup>

A major advantage of ZEB2 scoring is its nondependence on specialized molecular testing. Immunohistochemical analysis is routinely performed in all clinical pathology laboratories and readily translated into clinical practice. Most gene expression platforms quantify mRNA expression and transcriptional changes. However, an *in vitro* study<sup>42</sup> found that EMT is highly regulated by microRNAs; consequently, mRNA expression does not automatically result in protein expression and activation of EMT. A further limitation of polymerase chain reaction–based assays relates to tumor sampling; EMT often occurs in subsets of cancer cells within a tumor.<sup>43</sup> Therefore, attempts to quantify aberrant expression by sampling a small area may fail to represent true genetic heterogeneity of any sample. This limitation can be reduced by analysis of multiple sections of a single tumor by immunohistochemical analysis.

A key additional recognized feature of EMT is the acquisition of chemoresistance to compounds routinely used in clinical practice.<sup>9,44</sup> Therefore, it is possible that detection of a mesenchymal profile by ZEB2 expression not only identifies recurrence risk but also potential response to chemotherapeutic agents.

### Limitations

Although this study has demonstrated a consistent association among ZEB2 expression, early recurrence, and reduced survival in 3 independent cohorts, a few limitations need to be considered. Microsatellite stability, *kras*, and *BRAF* mutation status were not analyzed, and consequently, the association with ZEB2 was not studied. ZEB2 staining in the present study was based on surgical resection samples and not pretherapy biopsy samples. As a result, a small proportion of patients with CRC in both groups received neoadjuvant therapy, which may be a confounding factor. We did not note any association between ZEB2 status and neoadjuvant therapy and observed that even in the CRC cohort alone, ZEB2 was able to stratify patients; however, this study was not powered to address this subgroup formally. Finally, the association of the immune microenvironment with cancer outcomes is increasingly uncovered.<sup>45,46</sup> Similarly, there is increasing evidence that mesenchymal cancer cells can modulate the immune infiltrate to promote tumor progression.<sup>47</sup> Investigating the interdependence of immune scores and ZEB2 expression was beyond the scope of the current study; however, this may be critical in future efforts.

### Conclusions

We report, for the first time to our knowledge, that nuclear ZEB2 expression in CRC was associated with early recurrence and reduced survival in 2 independent prospective patient cohorts. ZEB2 identified patients at high risk of distant recurrence but not local recurrence, whereas subset analysis of node-negative disease revealed an identical association independent of stage. Addition of the ZEB2 score to nomograms composed of conventional TNM and histologic factors increased sensitivity and specificity of TNM, resulting in improved risk stratification in 14 of every 100 patients assessed.

Using independent testing, validation, and external cohorts and multivariable analysis, we found that ZEB2 expression in CRC may be associated with a high risk of distant recurrence after

curative surgery. Detection of this molecular marker by immunohistochemical analysis is a simple and reproducible method for detecting a mesenchymal phenotype. ZEB2 expression was an independent biomarker in all cases, including patients with node-negative and rectal cancer. Using a novel nomogram, we found that the addition of ZEB2 expression status more accurately estimated prognosis than did TNM staging alone and that TNM staging with ZEB2 expression status may have utility in patient counseling, trial design, and clinical decision making. Additional external validation or incorporation in a randomized biomarker-driven study are needed to support modification of TNM staging by addition of this novel marker.

---

## ARTICLE INFORMATION

**Accepted for Publication:** July 31, 2018.

**Published:** October 5, 2018. doi:10.1001/jamanetworkopen.2018.3115

**Open Access:** This is an open access article distributed under the terms of the [CC-BY License](#). © 2018 Sreekumar R et al. *JAMA Network Open*.

**Corresponding Author:** Alex H. Mirnezami, MD, PhD, Cancer Sciences Division, Faculty of Medicine, University of Southampton, Tremona Road, Somers Bldg, Southampton SO16 6YD, United Kingdom ([a.h.mirnezami@soton.ac.uk](mailto:a.h.mirnezami@soton.ac.uk)).

**Author Affiliations:** Cancer Sciences Division, Faculty of Medicine, University of Southampton, Southampton, United Kingdom (Sreekumar, Moutasim, Patel, Emo, Thomas, Primrose, Sayan, Mirnezami); Department of Surgery, Southampton University Hospital National Health Service Trust, Southampton, United Kingdom (Sreekumar, White, Primrose, Mirnezami); Medical Statistics and Mathematics Department, University of Southampton, Southampton, United Kingdom (Harris); Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts (DeMateos); Department of Molecular Biology and Genetics, Gebze Technical University, Gebze, Turkey (Yagci); Department of Cancer Sciences, University of Leicester, Leicester, United Kingdom (Tulchinsky).

**Author Contributions:** Drs Sayan and Mirnezami contributed equally to this work. Drs Sreekumar and Mirnezami had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

**Concept and design:** Sreekumar, Patel, Yagci, Tulchinsky, Primrose, Sayan, Mirnezami.

**Acquisition, analysis, or interpretation of data:** Sreekumar, Harris, Moutasim, DeMateos, Patel, Emo, White, Yagci, Thomas, Primrose, Sayan, Mirnezami.

**Drafting of the manuscript:** Sreekumar, Harris, Emo, White, Primrose, Sayan, Mirnezami.

**Critical revision of the manuscript for important intellectual content:** Harris, Moutasim, DeMateos, Patel, Yagci, Tulchinsky, Thomas, Primrose, Sayan, Mirnezami.

**Statistical analysis:** Sreekumar, Harris, Patel, White, Sayan.

**Obtained funding:** Sreekumar, Primrose, Sayan, Mirnezami.

**Administrative, technical, or material support:** Harris, Moutasim, Yagci, Tulchinsky, Primrose, Sayan, Mirnezami.

**Supervision:** Harris, Yagci, Thomas, Primrose, Sayan, Mirnezami.

**Conflict of Interest Disclosures:** None reported.

**Funding/Support:** Dr Sreekumar is supported by grant MR/L017539/1 from the Medical Research Council. Dr Sayan is supported by Innovation Grant T07 from Wessex Medical Research and grants from University of Southampton. Dr Mirnezami is supported by grant C28503/A10013 from Cancer Research UK/Royal College of Surgeons and Wessex Medical Research.

**Role of the Funder/Sponsor:** The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Additional Contributions:** The UK cancer research networks infrastructure and participating sites and staff and the Southampton Cancer Research UK Experimental Cancer Medicine Centre and Human Tissue Authority-approved tissue bank provided support to this study.



## REFERENCES

1. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. *Eur J Cancer*. 2010;46(4):765-781. doi:10.1016/j.ejca.2009.12.014
2. Manfredi S, Bouvier AM, Lepage C, Hatem C, Dancourt V, Faivre J. Incidence and patterns of recurrence after resection for cure of colonic cancer in a well defined population. *Br J Surg*. 2006;93(9):1115-1122. doi:10.1002/bjs.5349
3. Lugo TG, Braun S, Cote RJ, Pantel K, Rusch V. Detection and measurement of occult disease for the prognosis of solid tumors. *J Clin Oncol*. 2003;21(13):2609-2615. doi:10.1200/JCO.2003.01.153
4. Dienstmann R, Mason MJ, Sinicrope FA, et al. Prediction of overall survival in stage II and III colon cancer beyond TNM system: a retrospective, pooled biomarker study. *Ann Oncol*. 2017;28(5):1023-1031. doi:10.1093/annonc/mdx052
5. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015;21(11):1350-1356. doi:10.1038/nm.3967
6. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009;119(6):1420-1428. doi:10.1172/JCI39104
7. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol*. 2014;15(3):178-196. doi:10.1038/nrm3758
8. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008;133(4):704-715. doi:10.1016/j.cell.2008.03.027
9. Sayan AE, Griffiths TR, Pal R, et al. SIP1 protein protects cells from DNA damage-induced apoptosis and has independent prognostic value in bladder cancer. *Proc Natl Acad Sci U S A*. 2009;106(35):14884-14889. doi:10.1073/pnas.0902042106
10. Gupta PB, Onder TT, Jiang G, et al. Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell*. 2009;138(4):645-659. doi:10.1016/j.cell.2009.06.034
11. Kahlert C, Lahes S, Radhakrishnan P, et al. Overexpression of ZEB2 at the invasion front of colorectal cancer is an independent prognostic marker and regulates tumor invasion in vitro. *Clin Cancer Res*. 2011;17(24):7654-7663. doi:10.1158/1078-0432.CCR-10-2816
12. Shioiri M, Shida T, Koda K, et al. Slug expression is an independent prognostic parameter for poor survival in colorectal carcinoma patients. *Br J Cancer*. 2006;94(12):1816-1822. doi:10.1038/sj.bjc.6603193
13. Caramel J, Papadogeorgakis E, Hill L, et al. A switch in the expression of embryonic EMT-inducers drives the development of malignant melanoma. *Cancer Cell*. 2013;24(4):466-480. doi:10.1016/j.ccr.2013.08.018
14. Roepman P, Schlicker A, Tabernero J, et al. Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. *Int J Cancer*. 2014;134(3):552-562. doi:10.1002/ijc.28387
15. De Sousa E Melo F, Wang X, Jansen M, et al. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat Med*. 2013;19(5):614-618. doi:10.1038/nm.3174
16. Sadanandam A, Lyssiotis CA, Homicsko K, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med*. 2013;19(5):619-625. doi:10.1038/nm.3175
17. Marisa L, de Reyniès A, Duval A, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med*. 2013;10(5):e1001453. doi:10.1371/journal.pmed.1001453
18. Browne G, Sayan AE, Tulchinsky E. ZEB proteins link cell motility with cell cycle control and cell survival in cancer. *Cell Cycle*. 2010;9(5):886-891. doi:10.4161/cc.9.5.10839
19. Yamada S, Okumura N, Wei L, et al. Epithelial to mesenchymal transition is associated with shorter disease-free survival in hepatocellular carcinoma. *Ann Surg Oncol*. 2014;21(12):3882-3890. doi:10.1245/s10434-014-3779-2
20. Cellura D, Pickard K, Quarantino S, et al. miR-19-mediated inhibition of transglutaminase-2 leads to enhanced invasion and metastasis in colorectal cancer. *Mol Cancer Res*. 2015;13(7):1095-1105. doi:10.1158/1541-7786.MCR-14-0466
21. Bullock MD, Pickard K, Mitter R, et al. Stratifying risk of recurrence in stage II colorectal cancer using deregulated stromal and epithelial microRNAs. *Oncotarget*. 2015;6(9):7262-7279. doi:10.18632/oncotarget.3225
22. Zhang L, Pickard K, Jenei V, et al. miR-153 supports colorectal cancer progression via pleiotropic effects that enhance invasion and chemotherapeutic resistance. *Cancer Res*. 2013;73(21):6435-6447. doi:10.1158/0008-5472.CAN-12-3308

23. Ling H, Pickard K, Ivan C, et al. The clinical and biological significance of MIR-224 expression in colorectal cancer metastasis. *Gut*. 2016;65(6):977-989. doi:10.1136/gutjnl-2015-309372
24. Association of Coloproctology of Great Britain and Ireland. *Guidelines for the Management of Colorectal Cancer*. London, England: Association of Coloproctology of Great Britain and Ireland; 2007.
25. Oztas E, Avci ME, Ozcan A, Sayan AE, Tulchinsky E, Yagci T. Novel monoclonal antibodies detect Smad-interacting protein 1 (SIP1) in the cytoplasm of human cells from multiple tumor tissue arrays. *Exp Mol Pathol*. 2010;89(2):182-189. doi:10.1016/j.yexmp.2010.05.010
26. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. *PLoS Med*. 2012;9(5):e1001216. doi:10.1371/journal.pmed.1001216
27. Goswami CP, Nakshatri H. PROGeneV2: enhancements on the existing database. *BMC Cancer*. 2014;14:970. doi:10.1186/1471-2407-14-970
28. Krebs AM, Mitschke J, Lasierra Losada M, et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nat Cell Biol*. 2017;19(5):518-529. doi:10.1038/ncb3513
29. Ye X, Brabletz T, Kang Y, et al. Upholding a role for EMT in breast cancer metastasis. *Nature*. 2017;547(7661):E1-E3. doi:10.1038/nature22816
30. André T, Boni C, Mounedji-Boudiaf L, et al; Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) Investigators. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med*. 2004;350(23):2343-2351. doi:10.1056/NEJMoa032709
31. Kelley RK, Venook AP. Prognostic and predictive markers in stage II colon cancer: is there a role for gene expression profiling? *Clin Colorectal Cancer*. 2011;10(2):73-80. doi:10.1016/j.clcc.2011.03.001
32. Amado RG, Wolf M, Peeters M, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol*. 2008;26(10):1626-1634. doi:10.1200/JCO.2007.14.7116
33. Van Cutsem E, Köhne CH, Láng I, et al. Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. *J Clin Oncol*. 2011;29(15):2011-2019. doi:10.1200/JCO.2010.33.5091
34. Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell*. 2008;14(6):818-829. doi:10.1016/j.devcel.2008.05.009
35. Spaderna S, Schmalhofer O, Hlubek F, et al. A transient, EMT-linked loss of basement membranes indicates metastasis and poor survival in colorectal cancer. *Gastroenterology*. 2006;131(3):830-840. doi:10.1053/j.gastro.2006.06.016
36. Fujikawa H, Tanaka K, Toiyama Y, et al. High TrkB expression levels are associated with poor prognosis and EMT induction in colorectal cancer cells. *J Gastroenterol*. 2012;47(7):775-784. doi:10.1007/s00535-012-0532-0
37. Bellovin DI, Bates RC, Muzikansky A, Rimm DL, Mercurio AM. Altered localization of p120 catenin during epithelial to mesenchymal transition of colon carcinoma is prognostic for aggressive disease. *Cancer Res*. 2005;65(23):10938-10945. doi:10.1158/0008-5472.CAN-05-1947
38. Kroepil F, Fluegen G, Totikov Z, et al. Down-regulation of CDH1 is associated with expression of SNAI1 in colorectal adenomas. *PLoS One*. 2012;7(9):e46665. doi:10.1371/journal.pone.0046665
39. Busch EL, McGraw KA, Sandler RS. The potential for markers of epithelial-mesenchymal transition to improve colorectal cancer outcomes: a systematic review. *Cancer Epidemiol Biomarkers Prev*. 2014;23(7):1164-1175. doi:10.1158/1055-9965.EPI-14-0017
40. Puisieux A, Brabletz T, Caramel J. Oncogenic roles of EMT-inducing transcription factors. *Nat Cell Biol*. 2014;16(6):488-494. doi:10.1038/ncb2976
41. Vandewalle C, Comijn J, De Craene B, et al. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Res*. 2005;33(20):6566-6578. doi:10.1093/nar/gki965
42. Lamouille S, Subramanyam D, Billelo R, Derynck R. Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs. *Curr Opin Cell Biol*. 2013;25(2):200-207. doi:10.1016/j.ceb.2013.01.008
43. Cavallaro U, Christofori G. Cell adhesion in tumor invasion and metastasis: loss of the glue is not enough. *Biochim Biophys Acta*. 2001;1552(1):39-45.
44. Zheng X, Carstens JL, Kim J, et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature*. 2015;527(7579):525-530. doi:10.1038/nature16064

45. Pagès F, Mlecnik B, Marliot F, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet*. 2018;391(10135):2128-2139. doi:10.1016/S0140-6736(18)30789-X
46. Malietzis G, Giacometti M, Kennedy RH, Athanasiou T, Aziz O, Jenkins JT. The emerging role of neutrophil to lymphocyte ratio in determining colorectal cancer treatment outcomes: a systematic review and meta-analysis. *Ann Surg Oncol*. 2014;21(12):3938-3946. doi:10.1245/s10434-014-3815-2
47. Chen L, Gibbons DL, Goswami S, et al. Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. *Nat Commun*. 2014;5:5241. doi:10.1038/ncomms6241

**SUPPLEMENT.**

**eTable 1.** REMARK Biomarker Profile for the Present Study

**eTable 2.** TRIPOD Checklist: Prediction Model Development

**eTable 3.** Patient Demographics for Test and Validation Cohorts

**eTable 4.** Association Between Clinico-Pathological Features and ZEB2 Expression

**eTable 5.** Multi-variable Cox-regression analysis of overall and disease free survival in the test cohort (panel A) and validation cohorts (panel B), presented as hazard ratios (HR) with 95% confidence interval (CI).

**eTable 6.** Patient Demographics for Stage I and II Disease

**eTable 7.** Association Between Clinico-Pathological Features and ZEB2 Expression in Node-Negative Disease

**eTable 8.** Multi-variable Cox-regression analysis of overall and disease free survival in the node-negative disease, presented as hazard ratios (HR) with 95% confidence interval (CI)

**eFigure 1.** External validity was investigated using the open access portal PROGeneV2, which is compiled from GEO, EBI array express, and the The Cancer Genome Atlas.

**eFigure 2.** Nuclear ZEB2 associates with risk of distant (A) but not local (B) recurrence in test cohort.

**eFigure 3.** Nuclear ZEB2 associates with risk of distant (A) but not local (B) recurrence in validation cohort.

**eFigure 4.** ZEB2 is an Independent Prognostic Marker of Survival in Patients with Non-Metastatic Rectal Cancer (Sub-Group Analysis)